

SYSTEMS AND METHODS FOR APPLICATION SERVICE PROVISION

FIELD OF THE INVENTION:

The present invention relates to data processing and, more particularly, relates to systems
5 and methods for providing software applications and data processing to user communities over a
network in an efficient, low-overhead manner.

BACKGROUND OF THE INVENTION:

In recent years, there have been dramatic improvements in technologies that make
10 bandwidth available for data transmission. These improvements have resulted in ubiquitous
networks, such as the Internet, and have brought about rapid change in the operation of
numerous industries including the software industry.

Conventionally, the software industry has developed application software for proprietary
operating systems. Application software was then conventionally hosted on mainframe
15 computers with output from software applications provided to character based terminals pursuant
to proprietary protocols.

With the advent of inexpensive personal computers, this mainframe application software
delivery model changed to a client-server model in which application software developers
distributed application software programs to end users. In the latter scenario, the end users
20 loaded or downloaded the application software on their computer, a "fat client" machine, and ran
the application software directly on a proprietary operating system such as Microsoft Windows
or Unix.

Some of the application programs in a client-server model reside on a fat client and
require interaction with network resources, such as programs and data resident on servers within

the network. In other client – server models, the application programs reside on the server and are provided to the client system with the aid of emulation software on the client system.

With the advent of the Internet and the world-wide web, client systems have been implementing browser programs to present information received from a network to users. The browser programs include an application program interface (API) that programmers may use to create plug-ins that enable browsers to render previously unrecognized information, to recognize new communications protocols and to execute applications. Browser programs, supplemented with plug-ins as necessary, provide the flexibility to interact with software applications that are remotely executed on a network. Moreover, on the server side, an application program that was written for a proprietary operating system or display protocol may be web enabled and provided to browsers on remote client systems over a network. This entails translating the output from the application program into a protocol that is recognized by the browser program or an associated plug in.

The ability to web enable existing applications and remotely host them on a network provides advantages to application software vendors as well as end users of the software. Businesses called application service providers (ASPs) have arisen to facilitate providing application software to end users and their organizations over a network and, in some cases, to facilitate web enabling of software applications. ASP businesses allow users and their organizations the flexibility to rent, as opposed to purchase, software, to avoid time consuming installations of software on client systems and to order and use software on an as needed basis. For organizations, use of an ASP may effectively represent an outsourcing of maintenance operations and information services to the ASP. ASPs also allow software vendors additional software distribution channels from which to derive revenue from end users.

In order for ASP businesses to succeed in delivering software application service to end users, the ASP must be able to deliver reliable, high-performance, secure service that is convenient for organizations and users to configure. If any of these features is lacking, organizations and users may prefer local execution and control of the application software. ASP businesses must also confront problems of scalability, extensibility and integration. With respect to scalability, demand for application service for a particular ASP may exponentially increase several orders of magnitude over a short period of time. Therefore, scalability may be critical.

Accordingly, there is a need for an architecture and methods for providing application service that allow an ASP to commission new servers and equipment for delivering application service rapidly and without interrupting existing service. In addition, there is a need for robust architecture and methods that help prevent service disruption despite server and network link failures. There is a further need for an architecture and methods that make efficient use of server and other resources of the ASP in delivering service. There is still a further need for an architecture and methods that minimize administrative burdens associated with providing application service to organizations including, for example, burdens of providing users and organizations immediate and changeable access to applications and data associated with diverse proprietary operating systems, the ability to bill for service and to perform periodic data back-ups. There is still a further need for methods that maximize the value of the ASP architecture.

SUMMARY OF THE INVENTION:

According to the present invention, an architecture for providing software application service includes an intranet comprising redundant links to a network and redundant switches for reliable provision of application services to client systems over the network. The intranet

provides a common interface for managing organizations and their users, granting access to application software, including only certain versions thereof, and data sets, tracking usage of services and performing periodic backing up of data. The architecture of the intranet is scalable so that application, administrative and brokering servers may be quickly added to keep up with exponential increases in demand.

According to one embodiment of the invention, a method of efficiently provisioning application services for a plurality of diverse applications includes creating an organization entity within a data center, creating an organization unit for the organization entity and associating a group identification number with the organization entity. The method further includes propagating the organization unit and the group identification number for the organization entity to at least one application server within the data center. The method may further include collecting information about the organization entity and storing the collected information in an administrative database. The method may further include associating a suffix with the organization entity, verifying the uniqueness of the suffix within the data center and storing the suffix, the organization unit and the group identification number in an administrative database. Permission information for application services and data sets may also be stored in association with the organization entity in the administrative database.

The applications which form the basis of the application services may be published applications or custom applications. The applications may also be, for example, Windows based applications, Unix based applications, Linux based applications or other diverse applications. The organization information may be propagated to application servers within the data center based on an active directory or multi-master architecture.

The method may further include a facility for adding a user to the organization entity, associating a user identification with the user and propagating the user identification in association with at least one of the organization units and the group identification numbers to at least one application server within the data center. The user identification and associated permission information may be stored in the administrative database.

BRIEF DESCRIPTION OF THE FIGURES:

The above described features and advantages of the present invention will be more fully appreciated with reference to the detailed description and appended figures in which:

Fig. 1 depicts various client configurations for connecting to a data center from which application service provision services are provided according to embodiments of the present invention.

Fig. 2 depicts an embodiment of the architecture of a data center from which application service provision services are provided according to embodiments of the present invention.

Fig. 3 depicts an administrative server array according to an embodiment of the present invention..

Fig. 4 depicts a tarantella server array within an application service provider architecture according to an embodiment of the present invention.

Fig. 5 depicts Unix application server array within an application service provider architecture according to an embodiment of the present invention.

Fig. 6 depicts a windows application server array within an application service provider architecture according to an embodiment of the present invention.

Fig. 7 depicts a windows cluster server within an application service provider architecture according to an embodiment of the present invention.

Fig. 8 depicts a data storage unit within an application service provider architecture according to an embodiment of the present invention.

5 Fig. 9 depicts a method of defining organizations within a data center according to an embodiment of the present invention.

Fig. 10 depicts a method of adding users within a data center according to an embodiment of the present invention.

10 Fig. 11 depicts a functional view of a method of propagating organization and user data to a plurality of servers within a data center according to an embodiment of the present invention.

DETAILED DESCRIPTION:

15 According to the present invention, an architecture for providing software application service includes an intranet comprising redundant links to a network and redundant switches for reliable provision of application services to client systems over the network. The intranet provides a common interface for managing organizations and their users, granting access to application software, including only certain versions thereof, and data sets, tracking usage of services and performing periodic backing up of data. The architecture of the intranet is scalable
20 so that application, administrative and brokering servers may be quickly added to keep up with exponential increases in demand.

Fig. 1 depicts various client configurations for connecting to a data center from which application service provision services are provided according to embodiments of the present

invention. Referring to Fig. 1, a data center 100 is coupled to the client systems 120 via a network 110.

The network 110 may be a local area network, a wide area network, the public switched telephone network, the interconnected backbones, routers, bridges, switches and servers known as the Internet, other communications links and combinations thereof. The network may include direct electrical connections, wireless, optical or any other communications links, including analog, digital, circuit switched and packet switched, for transmitting information.

The client systems 120 may be general purpose computer systems which each incorporate modems or other communications technologies for exchanging information with the network 110. The client systems 120 may be coupled directly to the network 110 or may illustratively be coupled by way of a firewall 140, a proxy 150 or a LAN/WAN 160. Each client system may also be coupled to a printer or other peripherals 130. A printer or other peripheral 130 may also be coupled to the network 110 via a LAN/WAN 160 as shown.

Fig. 2 depicts an illustrative implementation of a data center for providing application services according to an embodiment of the present invention. Fundamentally, the architecture shown is flexible, robust and redundant. Referring to Fig. 2, the network 100 includes routers 200 coupled in parallel to the network 110. One of the routers 200 is within a left leg and the other is in a right leg of the network. The parallel connection is redundant to help prevent data center down time.

The routers 200 exchange packet data between the network 110 and the rest of the data center 200. The routers 200 receive and forward packets to appropriate elements within the data center 100 based on headers in the packets. The parallel switches 205 switch packets in the data center to steer packets in the appropriate direction. The switches 205 are interconnected as well

such that if a path in the direction of the left leg is broken, packets may be switched to the right leg.

The switches 205 are coupled to firewalls 210 in a criss-cross arrangement. Switches 215 are also coupled to the firewalls 210 in a criss-cross arrangement as shown. This arrangement permits packet traffic to by-pass one firewall 210 and travel through the other in the event of failure of one. In essence, the firewalls 210 look at each packet entering or leaving the network and accepts or rejects it based on user-defined rules. The firewall may apply application gateway techniques, circuit-level gateway techniques which apply certain tests prior to establishing a connection and/or proxy server techniques. Proxy server techniques effectively hide the true, internal data center network addresses from the network 110.

The switching routers 220 and 230 are coupled each coupled to the left and the right legs of the network and to each other. The switching routers 220 and 230 route data between and among a tarantella array 240, an administration array 245, a data storage unit 250 and a plurality of switches 225, 235 and 250. The switches 225 filter and forward packets between segments of the data center network. According to one embodiment of the invention, the data center network depicted is an ethernet network or a giga-bit ethernet network. The switches 225, 235 and 250 in this implementation may be used to implement a switched ethernet or giga-bit ethernet network.

The data storage unit 250 stores user application data for users of the ASP services. The data storage unit 250 serves files to the other functional units within the data center and to users at client systems 120 accessing the network.

Fig. 8 depicts an internal view of an embodiment of the data store. Referring to Fig. 8, the data storage unit includes two data movers 800 which provide access to a drive array. The drive array may comprise a redundant array of inexpensive drives (RAID) type storage device or

other storage device. The data movers 800 offer redundant access to the drive array 810 such that if one data mover 800 fails the other data mover 800 takes over. The data movers and array may be configured to provide storage in a network file system to allow users access to shared files stored in the array. There may be a separate system query language (SQL) path into the drive array 810 to facilitate database operations.

Fig. 3 depicts an administrative server array 245. The administrative server array 245 maintains data which identifies information for organizations and users of the data center and other details that are described below and propagates the data to the other functional components of the data center 100. The administrative server array 245 includes administrative servers 300 as shown. The administrative servers may each include an active directory 310 and an administrative database 320. The active directory 310 stores and automatically propagates administrative data to windows servers and other compatible servers. The administrative database 320 is used to store and propagate administrative data to UNIX based and compatible servers. The administrative server array and the servers themselves may be coupled to one of the switching routers 220 directly, however other convenient arrangements are possible. The population of the administrative database and the active directory to manage access control to the data center and other functions is described in more detail with reference to the method flow diagrams of Figs. 9-11.

Fig. 4 depicts an array of tarantella servers 400, which may be connected to the data center 100 network via the switching router 220. The tarantella array and servers within the array may be used as an intermediary between UNIX application servers within the data center and client systems coupled to the data center 100 via the network 110. Pursuant to this intermediary function, a client system which seeks to access a UNIX server does so via a

tarantella server. The tarantella server communicates with the client system according to a protocol called AIP and with the UNIX or other application according to a different protocol such as RDP. The client system includes corresponding capability to interface with the tarantella server pursuant to AIP and similarly the UNIX server communicates with the appropriate
5 Tarantella server according to the RDP or other compatible protocol. Tarantella servers and their functionality in brokering applications is set forth in U.S. Patent No. 6,104,392

Fig. 5 depicts a UNIX server array 260 which is coupled to the data center 100 network via a switching router 250. The UNIX server array includes a user accounts database 510 and an application data 520 portion. The application data portion stores data for users of the servers.
10 The application data may be physically resident in the data storage unit 250 or on the UNIX server itself. The UNIX servers store and execute application programs in response to authorized user requests to execute the applications. Access to the applications and application data is controlled by the user accounts.

During operation of the data center, a user may interact with browser software on the client system to access the data center. According to one embodiment of the invention, the user
15 be routed through the data center network to a tarantella server. The tarantella server may transmit an interactive web page back to the user which permits the user to launch applications, such as UNIX applications. When UNIX applications are launched in this manner, the user interacts with the Tarantella server via, for example, the AIP protocol. The AIP protocol delivers
20 to the client system display data and the application interface from the chosen UNIX application.

Fig. 6 depicts a windows server array which is analogous to the UNIX server array. However, the windows server array runs the Microsoft WindowsTM operating system. The windows server array may be coupled to the data center 100 network via the switches 250. The

windows server array includes an active directory 610 and an administrative database 620 for storing administrative information that may be used for application and file access control and other purposes. The windows server array also has application programs mounted on it with which users at client systems may interact via Tarantella as described above or via other protocols.

Fig. 7 depicts a Windows cluster server 270. The Windows cluster server 270 which may include cluster members 700. The cluster members 700 may be coupled together and to a shared data array 710. This arrangement provides another method for accessing the data storage via SQL.

Fig. 9 depicts a method of defining organizations within a data center according to an embodiment of the present invention. The method may be implemented by an administrative tool which amounts to a software program resident on the administrative server or another server for obtaining administrative information.

Referring to Fig. 9, in step 900 organization information is collected to define an organization to the data center 100. This information may include the name of the company, billing information, the name of a designated administrative contact and other information. According to one embodiment, this information includes a QORG suffix. The QORG (or QORG suffix) is a short name used to identify the organization and maintain user name uniqueness in the Data Center. (i.e. alx, m2m). The organization may be an individual or a corporation but in general is an organization or entity that is to be billed as a single unit. The organization may have associated with it a plurality of users that are entitled to ASP privileges with respect to particular applications. These users may be divided into various groups with various access privileges.

In step 910, an organization entity is created based on the information collected in step 900. In step 920, the uniqueness of the QORG suffix is checked by querying the administrative database to determine whether or not the QORG suffix is taken. If not, then another QORG must be chosen.

5 Step 930 may begin after QORG uniqueness is established. In step 930, a Windows organizational unit for the QORG is established. Then in step 940, a group identification (GID) number is associated with the QORG. In step 950, the information for the organization is stored into the administrative database. The information is also stored into the active directory.

10 In step 960, the GID is added to user data of the appropriate UNIX system and to the active directory of appropriate Windows systems. Organizations, for example, may be serviced by one or a subset of UNIX and Windows servers.

15 Fig. 10 depicts a method of adding users within a data center according to an embodiment of the present invention. The method of Fig. 10 is also be implemented by an administrative software tool. The administrative software tool may be run by on an administrative server and in general is also run pursuant to the ASP mode. Accordingly, designated administrators may interact directly with the administrative tool to define user access privileges and other privileges and features described below. This is powerful and allows a data center to roll out service to a large number of users with very little human capital required for administration because the organizations themselves perform, to a large extent, their own administration.

20 Referring to Fig. 10, in step 1000 a user (a designated administrator at a client system) is prompted by the administrative tool to take an action with respect to user administration. In step 1010 the tool determines whether the command is to add, modify or delete user data. IF the command is to ADD user data, then in step 1030, the tool receives user information including

permission information for applications, profiles, files and data. In step 1040, the user is added to an organizational unit within the user's QORG. In step 1050, a user identification (UID) number is associated with the user. Then in step 1060, the UID and user information is stored into the administrative database in association with the QORG. The information is also stored into the active directory. In step 1070, the user is added to the user accounts of appropriate UNIX systems to permit access to those systems. The systems are chosen based on the UID and GID of the user's organization. In this manner new users are identified to the data center and permitted to access ASP services and generate revenue for the data center. This may occur without any involvement by administrative personnel of the data center 100.

If in step 1010 the command is to modify a user, then step 1080 begins. In step 1080 user information may be received including permission information for applications, files profiles, and other information generally such as the user's name, address, phone number, email address, etc. In step 1090, the modified user information is stored into the administrative database in association with the QORG of the user. The modified user information is also stored into the active directory of Windows servers. In step 1095, the modified user id may be added to the user accounts of appropriate UNIX and Windows systems.

If in step 1010 the command is to delete a user then step 1020 begins. In step 1020, the UID and user information is deleted from the administrative database and active directories however a tombstone is saved.

Fig. 11 depicts a graphical illustration of the manner in which the administrative tool interacts with the administrative database, the active directory, the internal database of Tarantella servers and the user accounts of UNIX servers. When there is a change in organization or user information or administrative information generally, this information is propagated as illustrated.

5 modified or deleted information is then propagated to the user accounts and to the internal database of the Tarantella servers via a database merge program.

While particular embodiments have been disclosed, it will be understood by those having ordinary skill in the art that changes may be made to those embodiments without departing from the spirit and scope of the invention.